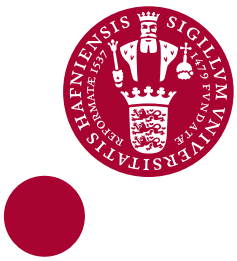


Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment

Pingel, Almut Maria; Hansen, Jørn Dines; Hansen, Jørgen Beck; Dam, Mogens; Xella, Stefania; Hansen, Peter Henrik; Petersen, Troels Christian; Thomsen, Lotte Ansgaard; Løvschall-Jensen, Ask Emil; Alonso Diaz, Alejandro; Monk, James William; Pedersen, Lars Egholm; Wiglesworth, Graig; Aad, G.; Galster, Gorm Aske Gram Krohn; Abbott, B.; Abdallah, J.; Abdinov, O.; Aben, R.; Abolins, M.; AbouZeid, O.S.; Abramowicz, H.; Abreu, H.; Abreu, R.

Published in:
Nuclear and Particle Physics Proceedings

DOI:
10.1016/j.nuclphysbps.2015.09.179

Publication date:
2016

Document version
Publisher's PDF, also known as Version of record

Available online at www.sciencedirect.com

ScienceDirect

Nuclear and Particle Physics Proceedings 273–275 (2016) 1141–1146

www.elsevier.com/locate/nppp

Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment

Almut Pingel, on behalf of the ATLAS Collaboration

Niels Bohr Institute, University of Copenhagen, Denmark

Abstract

This article gives an overview of the steps taken in ATLAS to identify hadronically decaying tau leptons and to validate the performance. The tau trigger, the reconstruction and identification algorithms, and the energy calibration are described. The performance is tested with $Z \to \tau\tau$ events, collected in 2012 at 8 TeV center-of-mass energy of the LHC. Identification efficiencies are determined both in real data and simulation and differences are expressed in terms of correction factors, with uncertainties below 6%. The uncertainty on the energy scale is measured with two independent methods and found to be less than 4%. All algorithms show good stability against a varying number of simultaneous proton-proton collisions.

Keywords: ATLAS, LHC, tau reconstruction, tau identification, tau trigger, efficiency measurement

1. Introduction

Tau leptons play an important role in the ATLAS [1] physics program at the LHC [2]. Examples are Standard Model analyses, ranging from cross section to polarization measurements [3, 4, 5, 6], and searches for physics beyond the Standard Model [7, 8, 9]. A very important milestone is the evidence for a Higgs boson decaying to a tau lepton pair [10].

This article describes the steps taken from the appearance of a tau lepton in the detector to the use in a physics analysis. First, the tau lepton is reconstructed and discriminated from other physics objects in the detector. Its energy is measured in the calorimeter with a calibration optimized for hadronically decaying tau leptons. All steps are validated using real data, mostly $Z \to \tau\tau$ events. At the end of this article, the tau lepton trigger is described briefly.

The tau lepton is the heaviest lepton and decays either leptonically, denoted τ_{lep}, or hadronically, denoted τ_{had}. A typical 50 GeV tau lepton travels ≈ 2 mm and decays before it even reaches the first layer of the ATLAS detector. It can therefore only be identified by its decay products. The electron or muon in a leptonic decay, $\tau \to e\nu_e\nu_\tau$ or $\tau \to \mu\nu_\mu\nu_\tau$, are nearly indistinguishable from prompt electrons or muons. The leptonic decay modes are therefore not considered in the τ identification algorithms. The hadronic mode, $\tau \to$ hadrons ν_τ, occurs in 65% of all cases. Predominantly, the visible part of the tau lepton decay is composed of one or three charged pions and zero to two neutral pions. This leads to a specific signature used to distinguish the hadronic τ decay from other objects in the detector.

2. Reconstruction and identification of τ_{had}

The reconstruction of hadronic τ decays [11] is initiated by anti-k_t jets [12] with a distance parameter of 0.4, taking calibrated TopoClusters [13] as input. The jets are required to fulfil $p_{\text{T}} > 10$ GeV and be within the region covered by the tracking systems, $|\eta| < 2.5$. The τ_{had} candidate is then built from TopoClusters within $\Delta R < 0.2$ around the jet center. The τ production vertex is identified before good quality tracks of at least 1 GeV are associated with the τ_{had} candidate within a cone of $\Delta R < 0.2$ around the τ_{had} axis. This results in a stable

http://dx.doi.org/10.1016/j.nuclphysbps.2015.09.179

reconstruction efficiency for a varying number of simultaneous proton-proton interactions per bunch crossing (*pile-up*).

2.1. QCD jet rejection

The main background for τ_{had} identification (*tau ID*) are jets initiated by gluons and quarks, called QCD jets from here on. The rate of QCD jet production is extremely high at the LHC and exceeds the tau lepton production by orders of magnitude. These jets are rejected by exploiting the specific τ_{had} characteristics: the collimation of the decay products, the low number of charged tracks and neutral clusters, and the slightly displaced decay vertex. These characteristics are parameterized in several variables, for instance the number of *isolation tracks* (counted in $0.2 < \Delta R < 0.4$ around the τ_{had} axis), the mean p_{T} weighted track distance and the fraction of energy deposited in the cone of $\Delta R < 0.1$ (Fig. 1).

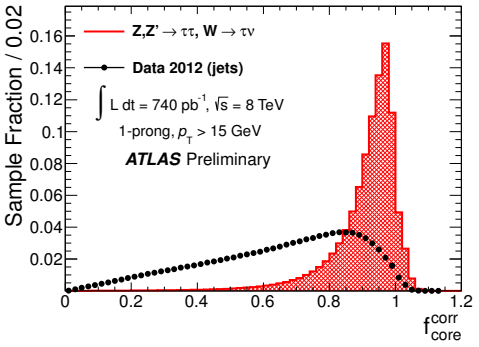
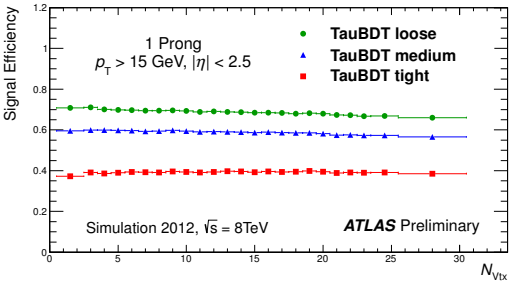


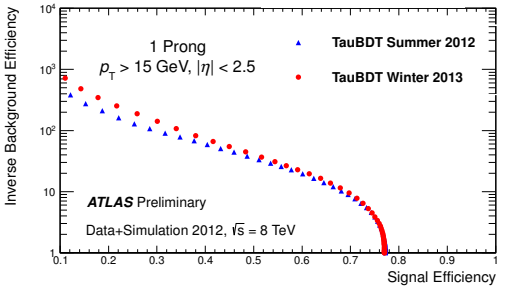
Figure 1: Discriminating variable $f_{\mathrm{core}}^{\mathrm{corr}}$: ratio of calorimeter energy in $\Delta R < 0.1$ to calorimeter energy in $\Delta R < 0.2$, for simulated hadronic τ decays (filled area) and multi-jet events in 2012 data (black points) with one associated track [11].

Since Winter 2012, a dedicated algorithm is used to reconstruct π^0 candidates within the hadronic τ decay, and the additional information is used for QCD jet discrimination. One example of such a new variable is the ratio of the momentum calculated with π^0 candidates and tracks (which represent the charged pions) to the momentum calculated using calorimeter clusters only.

The discriminating variables are combined in Boosted Decision Tree (BDT) classifiers. Three working points are pre-defined for different levels of signal efficiencies: 'loose', 'medium', and 'tight'. These are tuned to give flat efficiencies as a function of τ_{had} momentum. An issue of concern is the pile-up robustness of the tau ID algorithms. To decrease the pile-up dependence of the BDT classifier, the number of input variables was reduced and a pile-up correction is applied to calorimeter-based variables. The resulting classifier gives a signal efficiency which is flat as a function of the number of reconstructed vertices, as can be seen in Fig. 2(a) [14].



(a)



(b)

Figure 2: Performance of the algorithm for the rejection of QCD jets for τ_{had} candidates with one associated track: (a) signal efficiency for the three pre-defined working points as a function of the number of reconstructed vertices in the event and (b) inverse background efficiency as a function of signal efficiency without (circles) and with (triangles) additional π^0 information. The signal efficiencies are obtained from simulated $Z \to \tau\tau$, $Z' \to \tau\tau$ and $W \to \tau\nu$ events and are with respect to all true hadronic τ decays with one charged hadron. The background efficiencies are obtained from multi-jet events in data and are based on the number of jets being reconstructed as τ_{had} with one track [14].

Identification is provided for a minimum momentum of 15 GeV, but $p_{\mathrm{T}} \geq 20$ GeV is used in most physics analyses. This is due to the high background and higher uncertainties below this threshold. The rejection of QCD jets from the identification step alone ranges from a factor 10 to a few hundred (Fig. 2(b)) and is dependent on the momentum range and the number of tracks. Also the jet composition of the sample plays a role, because quark-dominated jets are more τ_{had}-like than gluon-dominated jets. The total rejection is considerably higher, as many QCD jets fail the requirement of a low core track multiplicity at the τ_{had} reconstruction step.

2.2. Lepton vetos

Also leptons, in particular electrons, are a background to τ_{had} identification. Besides the variables used for QCD jet rejection, information from the specific detector subsystems are helpful for the rejection of electrons. The Transition Radiation Tracker provides a powerful discriminator, because electrons are more likely to emit transition radiation. Also longitudinal shower information, such as the fraction of energy deposited in the electromagnetic calorimeter, is used. As for the QCD jet rejection, Boosted Decision Trees are trained and working points are defined to yield signal efficiencies of 75%, 85% and 95%. Slightly different variable sets are used in different pseudorapidity regions. The performance of the electron veto is shown in Fig. 3.

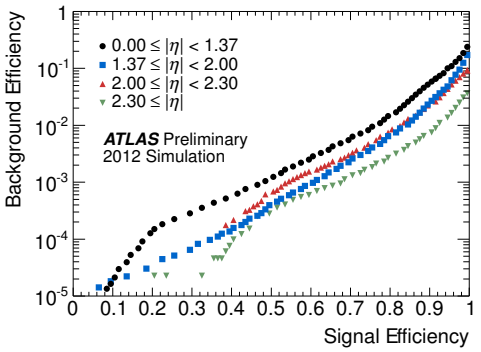
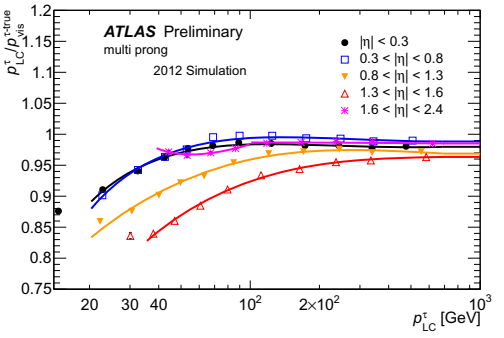
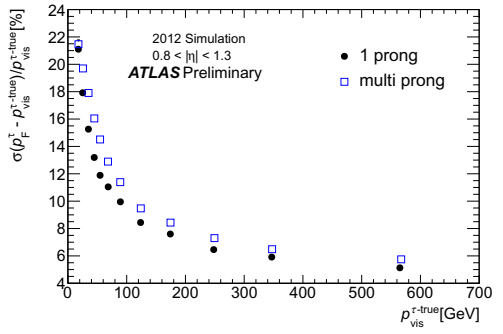


Figure 3: Background efficiency as a function of signal efficiency for the BDT based electron veto in different pseudorapidity regions. Signal efficiencies are obtained from simulated $Z \to \tau\tau$ events and background efficiencies from simulated $Z \to ee$ events [11].

Muons are unlikely to deposit enough energy in the calorimeter to be mistaken with hadronic τ decays and can generally be avoided by removing τ_{had} candidates that overlap geometrically with very loose muon candidates. A cut-based muon veto has been developed to reject muons that fail this overlap removal, because they fall into inefficient detector regions or are not reaching the muon system. These muons can be characterized by an unusually low or high electromagnetic energy fraction and track-momentum to calorimeter-energy ratio. As an example, very low momentum muons, that coincidentally overlap with calorimeter clusters, have a very low track-momentum to calorimeter-energy ratio, while muons that loose a significant fraction of their energy in the calorimeter, show the opposite behaviour. The muon veto has a signal efficiency better than 96% while rejecting 40% of the muons, and is depending both on the τ_{had} and muon identification working points.



(a)



(b)

Figure 4: (a) Response curve as a function of reconstructed momentum at LC scale, for multi-prong τ_{had} and different pseudorapidity regions and (b) momentum resolution as a function of true visible momentum for the region $0.8 < |\eta| < 1.3$ for 1-prong (circles) and multi-prong (squares) τ_{had}. 'Medium' identified τ_{had} from simulated $Z \to \tau\tau$, $Z' \to \tau\tau$ and $W \to \tau\nu$ events are used [15].

3. Tau energy scale

The energy of τ_{had} is estimated from the calorimeter energy deposits and is specifically calibrated [15]. First, the energy is brought to the so-called Local Calibration (LC) scale [16], which is applied to all jet objects and corrects for the non-compensation of the ATLAS calorimeter system, energy deposits outside the reconstructed clusters, and insensitive detector regions. The next calibration step takes into account specific decay and reconstruction characteristics, for instance that the energy is measured in a smaller cone for τ_{had} than for the seeding jets. This *tau energy scale* (TES) is reached in two steps. First, the ratio of the true visible energy and the reconstructed energy is obtained from simulated τ_{had} for different intervals of the true visible energy. The mean value in each interval in then evaluated as a function of the average energy at LC scale, and separately for different pseudorapidity regions and for decays with 1 track (*1-prong*) or more than 1 track (*multi-prong*), as

shown in Fig. 4(a). Furthermore, small corrections to account for inefficient detector regions and for pile-up contributions are applied.

The momentum resolution is shown in Fig. 4(b). It is around 20% for low momentum τ_{had} and saturates at around 5% for high momentum τ_{had}.

4. Performance measurement: Tau ID

The performance of the τ_{had} identification algorithms is validated with dedicated analyses on data [11]. Events are selected by a *tag* object and the τ_{had} algorithms are tested on a *probe* object. For the validation of the jet discrimination algorithms, samples enriched in $Z \to \tau\tau$ events are selected, where one of the tau leptons decays leptonically (tag) and the other hadronically (probe). Other leptonic Z and W events are suppressed by requirements on the missing transverse energy and the visible mass of the lepton and the τ_{had}. The remaining main background is composed of QCD jets and is estimated using a data-driven method. A template fit of the *extended track multiplicity* is performed. This variable was specifically developed for this measurement. It counts tracks in a wider radius around the τ_{had} in a pile-up robust way. Tracks with $p_{\text{T}} > 500\,\text{MeV}$ are added to the core tracks, if they are within $0.2 < \Delta R < 0.6$ and if there is at least one core track so that $p_{\text{T}}(\text{core track}) / p_{\text{T}}(\text{track}) \cdot \Delta R(\text{core track}, \text{track}) < 4.0$. Identification efficiencies can be obtained by fitting and extracting the number of τ_{had} with and without tau ID applied. The extended track multiplicity is shown in Fig. 5. The jet background is clearly reduced by applying tau ID and the lower distribution is dominated by real hadronic τ decays. By comparing the measured efficiencies in data and simulation, correction factors are obtained, which are applied in physics analyses to correct for the small mis-modeling in simulated samples. The data/MC correction factors are consistent with 1 for the 'loose' and 'medium' tau ID working points and around 0.9 for the 'tight' working point, with no dependence on p_{T} and only a small dependence on η. Uncertainties are of the order of (2-3)% for 1-prong τ_{had} and (4-6)% for multi-prong τ_{had}. Cross check analyses have been performed with $W \to \tau_{\text{had}}\nu_\tau$ and $t\bar{t}$ events. Consistent results are found in all channels, as shown in Fig. 6.

In the future it is planned to also provide *continuous* correction factors. Instead of measuring discrete factors for the pre-defined working points, the data to MC comparison is done on the BDT score directly, as shown in Fig. 7. This opens the possibility to explore the entire classifier in physics analyses or to choose the optimal signal efficiency individually in each analysis.

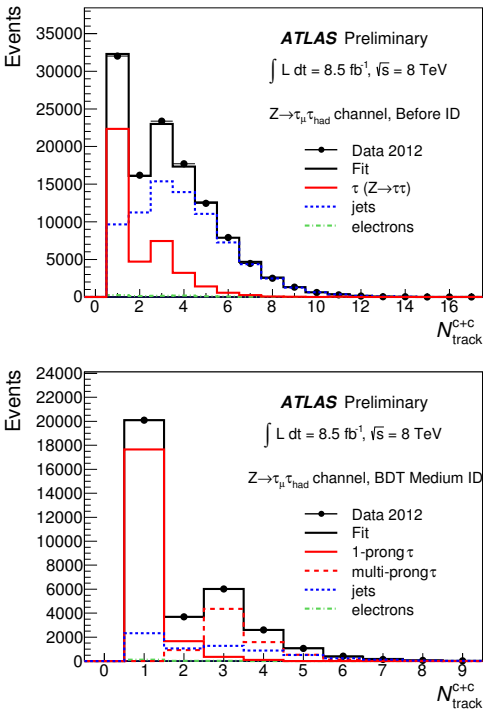


Figure 5: Extended track multiplicity after fit before tau ID (top) and after 'medium' tau ID (bottom). The τ_{had} signal template is taken from simulation, while the electron and jet templates are obtained from data in a control region [11].

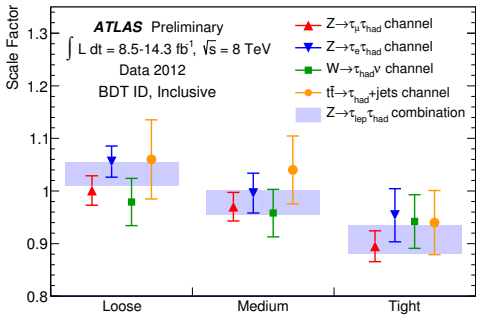


Figure 6: Summary of the data/MC correction (or scale) factors for the BDT tau ID, inclusive for 1-prong and multi-prong τ_{had}. The combined $Z \to \tau_{\text{lep}}\tau_{\text{had}}$ measurement (big rectangles) is used in physics analyses [11].

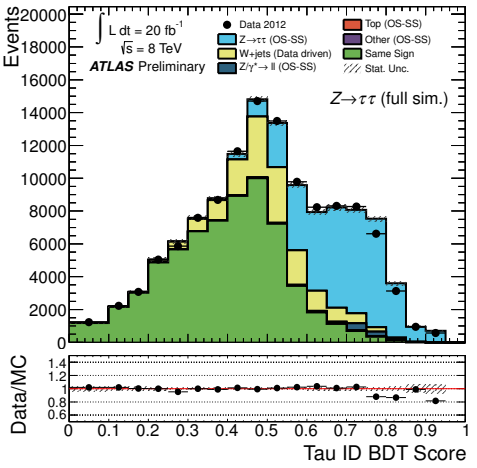


Figure 7: Tau ID BDT score for 1-prong τ_{had} selected with a $Z \to \tau\tau$ tag-and-probe analysis [14].

5. Performance measurement: TES

For the validation of the tau energy scale two independent approaches are used: a deconvolution method and a tag-and-probe measurement.

In the first method, the single particles involved in the hadronic τ decay are studied. The response of charged pions at low momenta is estimated with in-situ E/p measurements in low-pile-up data. At high momenta, test beam measurements are used for the central detector region and simulations otherwise. The neutral pion response is taken from studies of electrons from Z decays and minimum ionizing muons in the hadronic Tile calorimeter. Subsequently, these measurements are propagated to the full τ_{had} response, based on pseudo-experiments. This method gives access to the total tau energy scale uncertainty. It is estimated to be $\leq 3\%$ for τ_{had} with 1 track, and $\leq 4\%$ for τ_{had} with more than 1 track, over the full rapidity range for τ_{had} passing the 'medium' tau ID. The individual contributions to the uncertainty are shown in Fig. 8 for the most central η range.

The second approach is a tag-and-probe analysis using $Z \to \tau_\mu\tau_{\text{had}}$ events, with a similar event selection as in the tau ID performance study. The measurement is used as a cross-check of the first approach and is especially important to validate its simulation-based parts. A possible TES shift and its uncertainty is obtained by comparing the visible mass peak of the muon and the hadronic τ decay in data and simulation, as shown in Fig. 9. The measurement confirms the findings of the deconvolution method within uncertainty.

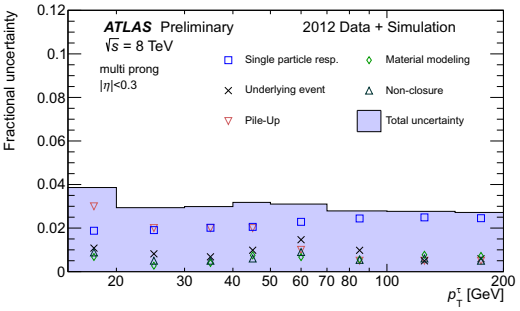


Figure 8: Uncertainty on the tau energy scale as a function of p_{T} for multi-prong τ_{had}. The individual components are shown as points, and the combined uncertainty as a blue filled band. p_{T} bins with equal uncertainties are grouped [15].

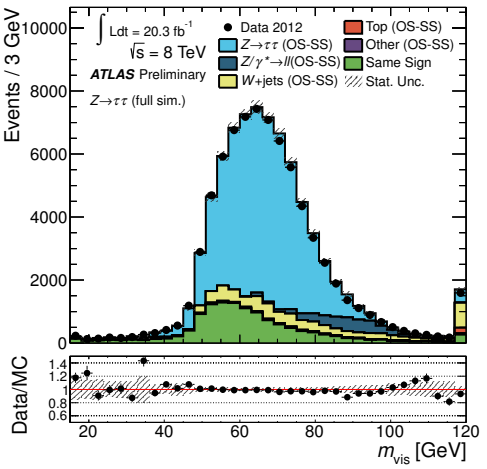


Figure 9: Visible invariant mass of the muon and the hadronic τ decay in $Z \to \tau_\mu\tau_{\text{had}}$ enriched events. The last bin is an overflow bin [14].

6. Tau lepton trigger

The τ_{had} reconstruction at the trigger level is special, as not the full detector information is available at all trigger stages and timing is much more critical. The ATLAS trigger consists of three stages: the first level is a hardware trigger. τ_{had} trigger objects are built using coarse calorimeter information and requiring calorimeter isolation to reduce the QCD jet contamination. The second trigger level is software based and has fast tracking and clustering available. Finer cuts on the calorimeter but also track isolation are used to select τ_{had} candidates. The algorithms applied at the last trigger stage are very similar to the offline tau ID. BDTs are used to separate hadronic τ decays from QCD jets, with variables mostly identical to those described in section 2.1.

Due to the high multi-jet background, single unprescaled τ_{had} triggers have high momentum thresholds.

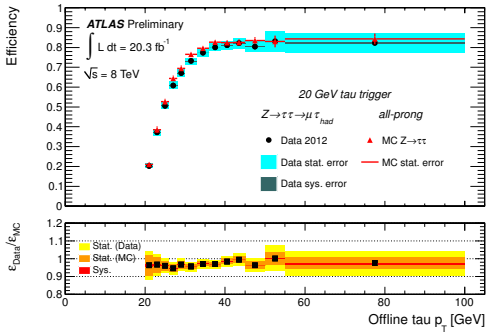


Figure 10: Trigger efficiency as a function of reconstructed p_{T}, as measured in data and simulation for 'medium' identified τ_{had} for the 20 GeV τ_{had} trigger. Expected background is subtracted from data. The uncertainty band on the ratio corresponds to statistical uncertainties in data and simulation and systematic uncertainties from the background subtraction [17].

To obtain a good sensitivity for physics processes with low momentum tau leptons, such as for the $H \to \tau\tau$ search, combined triggers are used. In pairing the single τ_{had} with a muon, an electron, a second τ_{had}, or missing transverse momentum, a τ_{had} trigger with a threshold of 20 GeV can be maintained. The trigger efficiencies and data/MC correction factors are obtained with a $Z \to \tau_\mu\tau_{\mathrm{had}}$ tag-and-probe measurement. Uncertainties are momentum dependent and of the order of (2-8)%, as shown in Fig. 10 [17].

7. Conclusion

The ATLAS τ_{had} identification methods based on Boosted Decision Trees are working well and provide good discrimination against jets and electrons. The good performance is demonstrated on data using $Z \to \tau\tau$ events, and cross-checked on $W \to \tau\nu$ and $t\bar{t}$ events. All measurements show consistent results. Data to MC correction factors are determined with uncertainties of (2-3)% for 1-prong τ_{had} and (4-6)% for multi-prong τ_{had}. The energy is measured in the calorimeter and calibrated specifically for hadronic τ decays. Uncertainties on the tau energy scale are $\leq 3\%$ for 1-prong τ_{had} and $\leq 4\%$ for multi-prong τ_{had}. By using combined triggers, a minimum τ_{had} trigger threshold of 20 GeV is maintained. The trigger efficiency is measured with uncertainties of (2-8)%, depending on momentum. Both identification and trigger algorithms show stability with varying pile-up conditions.

References

[1] ATLAS Collaboration, The ATLAS Experiment at the CERN Large Hadron Collider, JINST 3 (2008) S08003. doi:10.1088/1748-0221/3/08/S08003.

[2] L. Evans, P. Bryant (editors), LHC Machine, JINST 3 (2008) S08001. doi:10.1088/1748-0221/3/08/S08001.

[3] ATLAS Collaboration, Measurement of the W to $\tau\nu$ Cross Section in pp Collisions at $\sqrt{s}$ = 7 TeV with the ATLAS experiment, Phys. Lett. B706 (2012) 276-294. arXiv:1108.4101, doi:10.1016/j.physletb.2011.11.057.

[4] ATLAS Collaboration, Measurement of the Z to tau tau Cross Section with the ATLAS Detector, Phys. Rev. D84 (2011) 112006. arXiv:1108:2016, doi:10.1103/PhysRevD.84.112006.

[5] ATLAS Collaboration, Measurement of the top quark pair cross section with ATLAS in pp collisions at $\sqrt{s}$ = 7 TeV using final states with an electron or a muon and a hadronically decaying τ lepton, Phys. Lett. B717 (2012) 89-108. arXiv:1205.2067, doi:10.1016/j.physletb.2012.09.032.

[6] ATLAS Collaboration, Measurement of τ polarization in $W \to \tau\nu$ decays with the ATLAS detector in pp collisions at $\sqrt{s}$ = 7 TeV, Eur. Phys. J. C72 (2012) 2062. arXiv:1204.6720, doi:10.1140/epjc/s10052-012-2062-6.

[7] ATLAS Collaboration, Search for Supersymmetry in Events with Large Missing Transverse Momentum, Jets, and at Least One Tau Lepton in 21 fb^{-1} of $\sqrt{s}$ = 8 TeV Proton-Proton Collision Data with the ATLAS Detector, ATLAS-CONF-2013-026. http://cds.cern.ch/record/1525882.

[8] ATLAS Collaboration, A search for high-mass ditau resonances decaying in the fully hadronic final state in pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector, ATLAS-CONF-2013-066. http://cds.cern.ch/record/1562841.

[9] ATLAS Collaboration, Search for charged Higgs bosons in the τ+jets final state with pp collision data recorded at $\sqrt{s}$ = 8 TeV with the ATLAS experiment, ATLAS-CONF-2013-090. http://cds.cern.ch/record/1595533.

[10] ATLAS Collaboration, Evidence for Higgs Boson Decays to the $\tau^+\tau^-$ Final State with the ATLAS Detector, ATLAS-CONF-2013-108. http://cds.cern.ch/record/1632191.

[11] ATLAS Collaboration, Identification of the Hadronic Decays of Tau Leptons in 2012 Data with the ATLAS Detector, ATLAS-CONF-2013-064. http://cds.cern.ch/record/1562839.

[12] M. Cacciari, G. P. Salam, G. Soyez, The anti-kt jet clustering algorithm, JHEP 04 (2008) 063. arXiv:0802.1189, doi:10.1088/1126-6708/2008/04/063.

[13] W. Lampl et al., Calorimeter Clustering Algorithms: Description and Performance, ATL-LARG-PUB-2008-002. http://cds.cern.ch/record/1099735.

[14] ATLAS Collaboration, Tau Public Collision Plots, http://twiki.cern.ch/twiki/bin/view/AtlasPublic/TauPublicCollisionPlots, retrieved Sept. 8 2014.

[15] ATLAS Collaboration, Determination of the tau energy scale and the associated systematic uncertainty in proton-proton collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector at the LHC in 2012, ATLAS-CONF-2013-044. http://cds.cern.ch/record/1544036.

[16] T. Barillari et al., Local Hadronic Calibration, ATL-LARG-PUB-2009-001-2. http://cds.cern.ch/record/1112035.

[17] ATLAS Collaboration, Public Tau Trigger Plots for Collision Data, http://twiki.cern.ch/twiki/bin/view/AtlasPublic/TauTriggerPublicResults, retrieved Sept. 8 2014.